

# 사용자 정의 타입에 기반한 메타데이터 레지스트리를 위한 일관성 있는 질의 언어

김장원\*, 정동원\*\*, 백두권\*

\*고려대학교 컴퓨터·전파통신공학과

\*\*군산대학교 정보통계학과

e-mail: ikaros1223@korea.ac.kr

djeong@kunsan.ac.kr

baikdk@korea.ac.kr

## User-defined types Based Consistent Query Language for Metadata Registry

Jangwon Kim\*, Dongwon Jeong\*\*, Doo-Kwon Baik\*

\*Dept of Computer Science & Engineering, Korea University

\*\*Dept of Informatics & Statistics, Kunsan National University

### 요 약

이 논문에서는 메타데이터 레지스트리(ISO/IEC 11179: Metadata Registry)들이 가지고 있는 메타데이터 정보를 검색하고, 공유하기 위해 일관성 있는 질의 언어를 제안한다. 메타데이터 레지스트리는 국제 표준으로서 메타데이터들을 정의하고 이들을 관리 및 공유를 하기 위해 만들어졌다. 이와 같은 국제 표준을 기반으로 한 메타데이터 레지스트리들이 서지, 환경, 의료 분야 등의 다양한 영역에서 사용되고 있다. 이와 함께 메타데이터 레지스트리를 기반으로 하여 기존에 저장된 메타데이터들을 검색하고, 공유하고, 관리하고자 하는 이슈에 대한 연구가 진행되고 있다. 하지만 현재까지 다양한 분야에 있는 메타데이터 레지스트리가 가지고 있는 정보를 가져오기 위한 표준 인터페이스가 제공되고 있지 않다. 이러한 문제를 해결하기 위한 연구로 SQL에 메타데이터 레지스트리를 위한 공통 연산자들을 추가하여 메타데이터 레지스트리에 존재하는 데이터들을 활용하는 연구가 있다. 하지만 이런 연산자들을 이용하기 위해서는 상용 DBMS 엔진에 추가되어야 하며, 연산자들이 없는 경우 일관된 질의어를 수행할 수 없다는 문제를 가지고 있다. 따라서 이 논문에서는 이와 같은 문제를 해결하기 위해서 국제 표준인 SQL(ISO/IEC 9075)에서 정의하고 있는 사용자 정의 타입(User-defined types)을 기반으로 한 일관성 질의 언어를 제안한다.

키워드 : 메타데이터 레지스트리(MDR), 사용자 정의 타입(User-defined types), 일관성 질의(Consistent Query)

### 1. 서론

메타데이터 레지스트리는 메타데이터의 관리 및 상호운용성 향상을 위해 ISO/IEC JTC 1에 의해 개발된 국제 표준

으로 데이터의 표준화, 데이터의 정보 검색과 등록, 데이터에 대한 공유 등의 기능을 제공한다[1]. 이러한 장점 때문에 국내·외에서 다양한 분야에 있는 기관들이 메타데이터의 관리를 위해 ISO/IEC 11179 명세를 기반으로 하여 메타데이터 레지스트리를 구축해서 사용하고 있다[4,5]. ISO/IEC 11179에는 데이터 요소 등록, 관리, 수정 등에 대한 명세와 절차를 포함하고 있다. 이러한 메타데이터 레지스트리를 이

\* 이 연구에 참여한 연구자는 'BK 21 2단계 사업'의 지원을 받았으며, 이 논문은 2008년 정부(교육과학기술부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구(KRF-2008-314-D00485)인.

+ 이 저자는 이 논문의 교신저자임.

용하여 데이터를 공유 및 관리하는 연구가 진행되고 있다. 그 중의 하나로 다양한 메타데이터 레지스트리로부터 원하는 데이터를 가져올 수 있도록 하는 일관된 인터페이스를 제공하기 위한 연구가 있다(7). 이 연구에서는 메타데이터 레지스트리에서 필요한 연산자를 정의하여 표준 SQL을 확장하여 메타데이터 레지스트리로부터 원하는 정보를 가져온다. 하지만 이와 같은 경우 메타데이터 레지스트리에서 원하는 정보를 가져올 수 있도록 메타데이터 레지스트리를 위한 연산자가 사전에 정의 되어 있어야 한다. 그러므로 해당되는 연산자가 메타데이터 레지스트리에 존재 하지 않을 경우 연산을 수행할 수 없다. 그러므로 이러한 문제를 해결하기 위해서 본 논문에서는 SQL에서 정의하고 있는 사용자 정의 타입(User-defined types)를 이용하여 메타데이터 레지스트리에 원하는 연산을 할 수 있는 질의 언어를 제안한다.

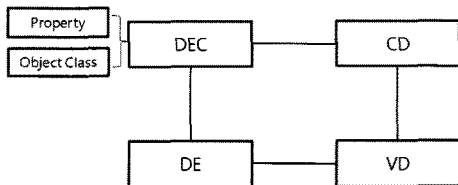
이 논문은 구성은 아래와 같다. II장에서는 관련 연구에 대해 기술하고, III장에서는 본 논문에서 제안하고 있는 질의 언어의 개념을 설명한다. IV장에서는 제안하는 SQL 질의 언어를 구현하여 실제 적용사례를 보여주며, 마지막으로 V장에서 결론 및 향후 연구에 대해 기술한다.

## II. 관련 연구

이 장에서는 메타데이터 레지스트리 및 기존 연구들의 문제점에 대해서 분석한다.

### 1. 메타데이터 레지스트리

ISO/IEC 11179는 국제 표준 기구인 ISO/IEC JTC1에 의해 개발되었다. 이는 메타데이터들의 공유 및 정보 교환을 할 수 있게 도와주며 일관성 있는 메타데이터 레지스트리를 구축하기 위한 국제 표준이다.



〈그림 1〉 메타데이터 레지스트리 메타모델: 주요 구성 요소

그림 1은 메타데이터 레지스트리의 주요 구성요소를 나타낸 것이다.

- 개념 영역(Conceptual Domain) : 데이터의 개념 정의
- 데이터 요소 개념(Data Element Concept) : 개념을 구체화 하기 위한 것으로 프로퍼티와 객체 클래스로 구성된다.
- 데이터 요소(Data Element) : 구체화된 데이터 요소, VD와 DEC를 기반으로 생성된다.
- 값 영역(Value Domain) : 데이터 타입, 다양한 표현 포함

- 프로퍼티(Property) : DEC에 대한 특성을 나타냄
- 객체 클래스(Object Class) : CD에 기반 한 클래스

### 2. SQL 확장을 통한 MDR 질의 언어

기존 연구로 메타데이터 레지스트리의 일관성 있는 접근을 위한 질의 언어가 있다. 이는 메타데이터 레지스트리에서 사용자가 원하는 메타데이터 정보를 얻기 위해서 사용자가 질의어를 통해 특정 메타데이터 정보를 얻고자 할 때, 그와 연관되어 공통적으로 사용하는 연산 패턴을 분석하고 이를 이용하여 표준 SQL을 확장한 질의 언어를 정의하고 설계하였다. 이를 통해 사용자들이 각기 다른 메타데이터 레지스트리들에서 원하는 정보를 얻어 올 때 동일한 인터페이스를 이용하여 메타데이터 레지스트리에 저장되어 있는 정보를 손쉽게 얻어 올 수 있었다. 하지만 사용자들이 메타데이터 레지스트리에서 연산자들을 사용하기 위해서는 기존의 상용 DBMS 엔진에 연산자들이 사전에 정의 되어 있어야 한다. 그렇지 않을 경우 SQL 구문 분석 및 의미 분석을 통해 메타데이터 레지스트리와 관련한 연산자를 구분해줘야 한다. 그러므로 연산자들이 들어가 있지 않는 각기 다른 상용 DBMS에서는 일관성 있는 연산을 할 수 없는 문제점이 발생한다.

### 3. 사용자 정의 타입 (User-defined types)

사용자 정의 타입은 SQL(ISO/IEC 9075) 명세에 정의 되어 있다(3). 사용자 정의 타입은 객체 타입의 하나로서 실제 세계에서 존재하는 엔티티들의 추상화된 객체이다(6). 이러한 객체 타입의 스키마 객체는 3가지 종류의 컴포넌트를 가진다.

- 1) 이름 : 데이터베이스 스키마에서 객체 타입을 식별하기 위해 사용한다.
- 2) 속성 : 실세계의 구조를 모델링한 것들로서 사용자 정의 타입이거나 다른 빌트인 타입일 수 있다.
- 3) 메서드 : 데이터베이스에 있는 속성 정보를 관리 및 수정하기 위해서 사용된다.

이를 이용하여 데이터베이스 내에 있는 기존의 테이블의 정보를 사용자가 원하는 형태로 만들 수 있다. 이를 통해 복잡한 질의어를 실행해 테이블간의 조인 및 연산을 할 필요 없이 사전에 원하는 값을 얻을 수 있다.

〈표 1〉 사용자 정의 타입과 메서드 사용 예

```

Create or replace TYPE EMP_TYPE AS OBJECT (
  EmpID integer,
  EmpName character varying(32),
  Title character varying(32),
  Salary integer,
  MEMBER FUNCTION showEmpName RETURN character
  varying
);
    
```

```

Create or eplace TYPE BODY EMP_TYPE AS
  MEMBER FUNCTION showEmpName RETURN character
  varying IS
  BEGIN
  RETURN EMPNAME;
  END;
END;
    
```

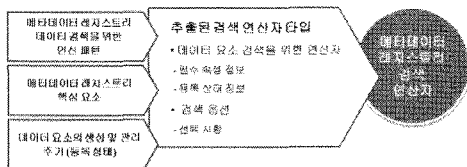
표 1은 Oracle DBMS에서 기본적으로 제공하는 사원 (Emp) 테이블을 이용하여 사용자 정의 타입을 생성한 예이다. 기존의 테이블과 테이블내의 속성들 중에서 필요한 것들만을 추출하여 사용자 정의 타입을 만들었다. 이러한 사용자 정의 타입이 포함된 SQL을 이용할 경우 사용자가 원하고자 하는 다양한 형태의 정보를 얻기 쉬워진다. 그러므로 이 테이블은 일종의 가공된 테이블로서 데이터 연산에 대한 효율성을 높여준다.

### III. 제안 질의 언어

이 절에서는 본 논문에서 제안하고 있는 질의 언어에 대한 접근 방법과 기본 정의에 대해서 기술한다.

#### 1. 접근 방법

이 절에서는 제안하는 질의 언어를 위한 접근 방법을 설명한다. 그림2는 메타데이터 레지스트리에서 원하는 정보를 얻기 위해 필요한 검색 연산자를 만드는 방법을 보여준다. 메타데이터 레지스트리를 위한 연산자는 사용자가 메타데이터 레지스트리를 통해 정보를 얻으려 할 때 필요로 하는 것들을 분석하여 연산 패턴을 추출하여 이를 반영해야 한다. 또한 메타데이터 레지스트리 내부의 핵심 요소들을 연산자를 통해서 가져올 수 있어야 한다. 그리고 메타데이터 레지스트리의 특징인 데이터 요소를 생성 및 관리하는 주기와 관련한 연산들도 포함해야 한다. 그러므로 연산자는 필수 속성 정보 및 등록 상태 정보 및 선택 사항을 추출할 수 있어야 한다. 이러한 기능을 수행하는 연산자를 위해 이 논문에서는 사용자 정의 타입을 정의하고, 메서드를 이용하여 필요한 연산을 수행할 수 있다.



〈그림 2〉 메타데이터 레지스트리 메타모델: 주요 구성 요소

이와 같은 접근 방법은 ISO/IEC 13249 Part3, 7에서도 사용되고 있다[2].

#### 2. 사용자 정의 타입 정의

사용자 정의 타입을 이용하여 메타데이터 레지스트리 표

준인 11179에서 정의하고 있는 메타모델들을 위한 연산자를 정의한다. 이 때 사용자 정의 타입은 메타데이터 레지스트리에서 포함하고 있는 메타 모델들 전부를 포함한다. 또한 각각의 메타 모델들이 가지고 있는 관계와 속성 정보를 포함한다. 사용자 정의 타입을 다음과 순서로 생성한다.

- 1) 메타데이터 레지스트리의 메타 모델 추출
  - 메타 모델 정보 추출
  - 메타 모델들의 관계 추출
  - 메타 모델의 속성 추출
- 2) 추출된 메타 모델 정보를 기반으로 한 사용자 정의 타입 정의
  - 각 메타 모델이 포함하고 있는 속성 정의
  - 속성을 검색하고 활용하기 위한 메서드 정의

다음은 정의된 사용자 정의 타입을 보여준다.

- Administration\_record\_type : 관리 정보
- Registration\_authority\_type : 등록 권한 정보
- Submission\_type : 제출 정보
- Stewardship\_type : 관리자 정보
- Context\_type : 문맥(상황) 정보
- Reference\_Document\_type : 참조 문서 정보
- Conceptual\_Domain\_type : 개념 영역 정보
- Data\_Element\_Concept\_type : 데이터 요소 개념 정보
- Data\_Element\_type : 데이터 요소 정보
- Value\_Domain\_type : 값 영역 정보
- Object\_Class\_type : 객체 클래스 정보
- Representation\_Class\_type : 표현 클래스 정보

#### 3. 사용자 정의 타입 메서드 정의

사용자 정의 타입을 이용하여 메타모델 정보를 특정 타입으로 생성을 하면 메타모델이 포함하고 있는 속성정보를 메서드를 이용해서 접근할 수 있으며 활용 할 수 있다. 다음은 메타데이터 레지스트리의 메타모델의 하나인 데이터 요소가 가지고 있는 속성들을 활용하기 위해 정의한 메서드들이다. 각각의 사용자 정의 타입 안에 존재하는 속성들에 대한 정보는 메서드를 호출함으로써 사용할 수 있다. 또한 속성을 대상으로 한 삽입, 삭제, 업데이트 연산을 수행하고자 할 경우 메서드에 추가적인 파라미터를 이용하여 원하는 작업을 수행할 수 있다.

〈표 2〉 데이터 요소를 위한 메서드

```

Data_element_administration_record()
Item_regeistration_authority_identifier()
International_code_designator()
Organization_identifier()
DATA_identifier()
...
Version()
    
```

사용자 정의 타입은 표준 SQL 명세에 포함되어 있으므로, 메타데이터 레지스트리에 본 논문에서 제안한 사용자 정의 타입을 기반으로 한 테이블 생성을 만들 수 있다. 사용자 정의 타입을 기반으로 만들어진 테이블을 서로 동일한 형태를 가지며, 동일한 형태의 메시지를 가지게 됨으로써 이를 통해 각기 다른 메타데이터 레지스트리에 대해 일관성 있는 질의를 수행할 수 있으며, 이를 통해 사용자가 원하는 정보를 가져 올 수 있다. 또한, 메타데이터 레지스트리가 상용 DBMS 위에 구축되었을 경우 기존 관련 연구에서 살펴봤던 연산자를 따로 구현해 줄 필요 없기 때문에 메타데이터 레지스트리를 위한 연산자를 데이터베이스 엔진에 추가적으로 개발하기 위한 비용이 들지 않게 된다.

#### IV. 구현 및 평가

이 장에서는 사용자 정의 타입을 기반 한 질의어를 상용 DBMS에 적용하여 생성한 예를 보이고, 평가 결과에 대하여 기술한다.

##### 1. 사용자 정의 타입과 메서드 구현

표 3은 메타데이터 레지스트리에 존재하는 메타 모델을 기반으로 사용자 정의 타입과 메서드를 보여준다. 메타데이터 레지스트리의 메타 모델중의 하나인 데이터 요소 중 일부 속성 정보를 기반으로 만든 사용자 정의 타입이다.

〈표 3〉 데이터 요소를 위한 사용자 정의 타입과 메서드 정의

```
CREATE TYPE Data_Element As (
  Data_element_administration_record NUMBER,
  Item_registration_authority_identifier NUMBER,
  Organization_part_identifier NUMBER,
  OPI_source VARCHAR2(2000),
  ....
  DATA_identifier NUMBER );
METHOD Data_element_administration_record()
  RETURNS NUMBER,
METHOD Item_registration_authority_identifier()
  RETURNS NUMBER,
METHOD International_code_designator()
  RETURNS VARCHAR2,
...
METHOD Organization_part_identifier()
  RETURNS NUMBER
```

표 3에서 사용자 정의 타입에 들어 있는 속성들은 메타데이터 레지스트리의 메타 모델중의 하나인 데이터 요소에서 정의하고 있는 속성들과 일치한다. 이러한 속성들의 정보를 접근하고, 업데이트와 같은 연산을 수행하기 위해서 메서드가 필요하다. 이와 같은 방식으로 메타데이터 레지스트리에 존재하는 메타모델들의 정보를 사용자 정의 타입들을 이용해 생성한다. 이렇게 정의된 사용자 정의 타입 및 메서드를 이용하여 다양한 메타데이터 레지스트리 간의 정보를 접근, 검색

및 관리함에 있어 사용자에게 질의어의 일관성을 제공한다.

#### 2. 평가

본 논문에서 제안하는 질의 언어를 연산자를 기본으로 한 SQL확장 연구와 비교하면, 연산자 생성 및 분석 및 추출을 위해 필요한 비용을 줄일 수 있다는 장점을 가진다. 또한, 표준 SQL에서 제공하는 사용자 정의 타입과 메서드를 이용함으로써 다양한 상용 DBMS에 구축된 메타데이터 레지스트리에 대해서도 일관성 있는 질의를 수행할 수 있다.

#### V. 결론

이 논문에서는 메타데이터 레지스트리를 접근하기 위한 기존의 방법들과 문제점을 정의하고 보다 효율적이며, 편리한 질의 언어를 제안하였다. 일관성 있는 질의 언어를 개발하기 위해 표준 SQL에서 정의하고 있는 사용자 정의 타입과 메서드를 사용하였다. 사용자 정의 타입은 ISO/IEC 11179 메타데이터 레지스트리의 명세에서 포함하고 있는 메타모델을 반영하였으며, 사용자는 메서드를 호출하여 사용자 정의 타입에 포함되어 있는 속성 정보들을 얻어오거나, 업데이트 할 수 있었다. 이를 통해 메타데이터 레지스트리에 사용자가 원하는 연산을 수행할 수 있었으며, 다양한 메타데이터 레지스트리에 존재하는 데이터들을 일관성 있는 인터페이스를 통해 얻어 올 수 있음을 확인할 수 있다.

#### 참고문헌

- [1] ISO/IEC JTC 1/SC 32/WG 2, "ISO/IEC 11179: Specification and standardization of data elements, Part 1~6", 2003.
- [2] ISO/IEC JTC 1/SC 32, "ISO/IEC 13249: Information technology-Database languages-SQL Multimedia and Application Packages", 2003
- [3] ISO/IEC JTC 1/SC 32/WG 3, "ISO/IEC 9075, Database Language SQL3, Part1~10", 1999.
- [4] KISTI, A study on the development of standardization and management model for science and technology information, Research Report, 2002.
- [5] EPA, Environmental Data Registry, <http://www.epa.gov/edr/>.
- [6] Stonebraker, M. "Architectual Options for Object-Relational DBMSs", White Paper, Informix Software, Inc., 1997
- [7] 신동길, 정동원, 김영갑, 박수현, 백두권, "메타데이터 레지스트리의 일관성 있는 접근을 위한 질의 언어," 한국정보과학회, 정보과학회논문지: 데이터베이스, 제 31권, 제 6호, pp.609-623, 2004년 12월